SEQUENCE LISTING

- (1) GENERAL INFORMATION:
 - (i) APPLICANT: SOPPET, DANIEL R. RUBEN, STEVEN M.
 - (ii) TITLE OF INVENTION: HEMATOPOIETIC SIGNALING FACTOR
 - (iii) NUMBER OF SEQUENCES: 15
 - (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
 - (B) STREET: 1100 NEW YORK AVENUE, SUITE 600
 - (C) CITY: WASHINGTON
 - (D) STATE: DC
 - (E) COUNTRY: US
 - (F) ZIP: 20005-3934
 - (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
 - (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: To be assigned
 - (B) FILING DATE: Herewith
 - (C) CLASSIFICATION:
 - (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 60/035,577
 - (B) FILING DATE: 16-JAN-1997
 - (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: STEFFE, ERIC K
 - (B) REGISTRATION NUMBER: 36,688
 - (C) REFERENCE/DOCKET NUMBER: 1488.0600001
 - (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (202) 371-2600
 - (B) TELEFAX: (202) 371-2540
- (2) INFORMATION FOR SEQ ID NO:1:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1545 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: cDNA
 - (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 66..1202
 - (ix) FEATURE:

(A) NAME/KEY: mat_peptide (B) LOCATION: 144..1202

(ix) FEATURE:

(A) NAME/KEY: sig_peptide
(B) LOCATION: 66..141

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CGGGCTTCGG GTCGGTGCAA GGCAGGCGCA CGGGGAAGGG CGCGCCGCG GGCCGCCACC 60						
CCACC ATG CTC AAG CGC TGC GGC CGG CGC CTG CTG CTG GCG CTG GCG Met Leu Lys Arg Cys Gly Arg Arg Leu Leu Ala Leu Ala -26 -25 -20 -15	107					
GGC GCG CTG CTC GCC TGC CTG GTG CTC ACC GCC GAC CCG CCG Gly Ala Leu Leu Ala Cys Leu Leu Val Leu Thr Ala Asp Pro Pro -10 -5 1	155					
CCT CCA CTG CCC GCC GAG CGC GGC CGG CGC GCG CTG CGC AGC CTG GCG Pro Pro Leu Pro Ala Glu Arg Gly Arg Arg Ala Leu Arg Ser Leu Ala 5 10 15 20	203					
GGC CCC GCG GGG GCT GCC CCG GCG CCC GGG CTG GGG GCG GC	251					
GCG CCC GGG GCG CTG GTC CGC GAC GTG CAC AGT CTG TCC GAG TAC TTC Ala Pro Gly Ala Leu Val Arg Asp Val His Ser Leu Ser Glu Tyr Phe 40 45 50	299					
AGC CTG CTC ACC CGC GCG CGC AGA GAT GCG GGC CCG CCC GGG GCT Ser Leu Leu Thr Arg Ala Arg Arg Asp Ala Gly Pro Pro Pro Gly Ala 55 60 65	347					
GCC CCC CGC CCC GCC GAC GGC CAC CCG CGC CCC CTG GCC GAG CCG CTC Ala Pro Arg Pro Ala Asp Gly His Pro Arg Pro Leu Ala Glu Pro Leu 70 75 80	395					
GCG CCC CGA GAC GTC TTC ATC GCT GTC AAG ACC ACC AAA AAG TTC CAC Ala Pro Arg Asp Val Phe Ile Ala Val Lys Thr Thr Lys Lys Phe His 90 95 100	443					
CGC GCG CGC CTC GAC CTG CTG CTG GAG ACC TGG ATC TCG CGC CAC AAG Arg Ala Arg Leu Asp Leu Leu Leu Glu Thr Trp Ile Ser Arg His Lys 105 110 115	491					
GAG ATG ACG TTC ATC TTC ACT GAC GGG GAA GAT GAG GCC CTG GCC AGG Glu Met Thr Phe Ile Phe Thr Asp Gly Glu Asp Glu Ala Leu Ala Arg 120 125 130	539					
CAC ACG GGC AAC GTG GTC ATC ACA AAC TGC TCG GCC GCC CAC AGC CGC His Thr Gly Asn Val Val Ile Thr Asn Cys Ser Ala Ala His Ser Arg 135 140 145	587					
CAG GCG CTG TCC TGC AAG ATG GCC GTG GAG TAT GAC CGC TTC ATC GAG Gln Ala Leu Ser Cys Lys Met Ala Val Glu Tyr Asp Arg Phe Ile Glu 150 155 160	635					

TCC GGC AG Ser Gly Ar 165					Asp Asn			683
CTG CGG AC Leu Arg Th		Arg Leu					g Asp	731
GTC TAC GT Val Tyr Va				Arg Pro				779
CGG GTC AG Arg Val Se 21	r Glu Asn					Phe Al		827
GGC GGC GC Gly Gly Al 230								875
CCG TGG GC Pro Trp Al 245					Ala Glu			923
CTG CCT GA Leu Pro As		Thr Ile					u Gly	971
GTG CCC CT Val Pro Le				His Ser				1019
CAG CAG GT Gln Gln Va 29	l Pro Thr					Leu Se		1067
GGT ATG TT Gly Met Ph 310								1115
TCG GTG GA Ser Val Gl 325					Ile His			1163
TAC CCG GA		Trp Cys					CCAT	1212
GGCTGAGACC	CAATCCCT	GG GCGCC	CCTGG TA	TCCAAAGO	GCCCAGG	GAC CCI	GTTGCGC	1272
TGCCCTGGCC	TCGGCATT	CG AGGCT	CCCCT AG	GGCCGTG	CTGTGCG	TGT GCG	TGTGCGT	1332
GTGTGTGTGT	GTGTACTG	CA TGCCC	ACCCG GG	TAGCAGG	TGCTGGG	CAG TTC	TGCTCTG	1392
TGGAGGGGCG	GGCACCAG	CG CCACT	TATGT GC	CTCTGCT	: CGAGGGC	CAG TGG	TATGGAG	1452
GGTCTGCTTG	GAGGAAGG	AT TTGTG	TGTCG GA	GGACACTO	: CGAGGGC	AAT TCI	'GTTAGGA	1512
TTTGTGGATC TTTCTACAGC TACGGGGCTC CGG								1545

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 379 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Leu Lys Arg Cys Gly Arg Arg Leu Leu Leu Ala Leu Ala Gly Ala
-26 -25 -20 -15

Leu Leu Ala Cys Leu Leu Val Leu Thr Ala Asp Pro Pro Pro Pro -10 -5 5

Leu Pro Ala Glu Arg Gly Arg Arg Ala Leu Arg Ser Leu Ala Gly Pro
10 15 20

Ala Gly Ala Ala Pro Ala Pro Gly Leu Gly Ala Ala Ala Ala Pro 25 30 35

Gly Ala Leu Val Arg Asp Val His Ser Leu Ser Glu Tyr Phe Ser Leu
40 45 50

Leu Thr Arg Ala Arg Arg Asp Ala Gly Pro Pro Pro Gly Ala Ala Pro 55 60 65 70

Arg Pro Ala Asp Gly His Pro Arg Pro Leu Ala Glu Pro Leu Ala Pro
75 80 85

Arg Asp Val Phe Ile Ala Val Lys Thr Thr Lys Lys Phe His Arg Ala 90 95 100

Arg Leu Asp Leu Leu Glu Thr Trp Ile Ser Arg His Lys Glu Met 105 110 115

Thr Phe Ile Phe Thr Asp Gly Glu Asp Glu Ala Leu Ala Arg His Thr 120 125 130

Gly Asn Val Val Ile Thr Asn Cys Ser Ala Ala His Ser Arg Gln Ala 135 140 145 150

Leu Ser Cys Lys Met Ala Val Glu Tyr Asp Arg Phe Ile Glu Ser Gly
155 160 165

Arg Lys Trp Phe Cys His Val Asp Asp Asp Asn Tyr Val Asn Leu Arg 170 175 180

Thr Leu Leu Arg Leu Leu Ala Ser Tyr Pro His Thr Arg Asp Val Tyr 185 190 195

Val Gly Lys Pro Ser Leu Asp Arg Pro Ile Gln Ala Met Glu Arg Val 200 205 210

Ser Glu Asn Lys Val Arg Pro Val His Phe Trp Phe Ala Thr Gly Gly 225 230

- Ala Gly Phe Cys Ile Ser Arg Gly Leu Ala Leu Lys Met Ser Pro Trp 235 240 245
- Ala Ser Gly Gly His Phe Met Asn Thr Ala Glu Arg Ile Arg Leu Pro 250 255 260
- Asp Asp Cys Thr Ile Gly Tyr Ile Val Glu Ala Leu Leu Gly Val Pro 265 270 275
- Leu Ile Arg Ser Gly Leu Phe His Ser His Leu Glu Asn Leu Gln Gln 280 285 290
- Val Pro Thr Ser Glu Leu His Glu Gln Val Thr Leu Ser Tyr Gly Met 295 300 305 310
- Phe Glu Asn Lys Arg Asn Ala Val His Val Lys Gly Pro Phe Ser Val 315 320 325
- Glu Ala Asp Pro Ser Arg Phe Arg Ser Ile His Cys His Leu Tyr Pro 330 335 340
- Asp Thr Pro Trp Cys Pro Arg Thr Ala Ile Phe 345 350
- (2) INFORMATION FOR SEQ ID NO:3:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 375 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: protein
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:
 - Met Leu Lys Asn Trp Gly Lys Lys Leu Leu Leu Ser Ile Val Gly Ala 1 5 10 15
 - Thr Leu Thr Cys Leu Leu Val Leu Val Val Asp Gln Gln Ser Arg His 20 25 30
 - Met Leu Glu Thr Gln Ser Asp His Glu Pro Gly Ser Ala Ala Val 35 40 45
 - His Leu Arg Ala Asp Leu Asp Pro Ala Asn Pro Gly Asp Gly Gly Asp 50 55 60
 - Pro Ala Asn Ser Ala Gln Asp Ser Gly Thr Phe Ser Ala Tyr Phe Asn 65 70 75 80
 - Lys Leu Thr Arg Val Arg Arg Asp Val Glu Gln Val Ala Ala Pro Ser 85 90 95
 - Lys Asp Ser Ala Ala Pro Glu Glu Asp Ile Thr Ala Asn Asp Val Phe
 100 105 110

Ile Ala Val Lys Thr Thr Lys Lys Phe His Arg Ser Arg Met Asp Leu 115 120 125

Leu Met Asp Thr Trp Ile Ser Arg Asn Lys Glu Gln Thr Phe Ile Phe 130 135 140

Thr Asp Gly Glu Asp Glu Glu Leu Gln Lys Lys Thr Gly Asn Val Glu 145 150 155 160

Ser Thr Asn Cys Ser Ala Ala His Ser Arg Gln Ala Leu Ser Cys Lys 165 170 175

Met Ala Val Glu Tyr Asp Lys Phe Ile Glu Ser Asp Lys Lys Trp Phe 180 185 190

Cys His Val Asp Asp Asp Asn Tyr Val Asn Val Arg Thr Leu Val Lys 195 200 205

Leu Leu Ser Arg Tyr Ser His Thr Asn Asp Ile Tyr Ile Gly Lys Pro 210 215 220

Ser Leu Asp Arg Pro Ile Gln Ala Thr Glu Arg Ile Ser Glu Ser Asn 225 230 235 240

Met Arg Pro Val Asn Phe Trp Phe Ala Thr Gly Gly Ala Gly Phe Cys 245 250 255

Ile Ser Arg Gly Leu Ala Leu Lys Met Ser Pro Trp Ala Ser Gly Gly 260 265 270

His Phe Met Asn Thr Ala Glu Lys Ile Arg Leu Pro Asp Asp Cys Thr 275 280 285

Ile Gly Tyr Ile Ile Glu Ser Val Leu Gly Val Lys Leu Ile Arg Ser 290 295 300

Asn Leu Phe His Ser His Leu Glu Asn Leu His Gln Val Pro Gln Ser 305 310 315 320

Glu Ile His Asn Gln Val Thr Leu Ser Tyr Gly Met Phe Glu Asn Lys 325 330 335

Arg Asn Ala Ile Leu Met Lys Gly Ala Phe Ser Val Glu Glu Asp Pro 340 345 350

Ser Arg Phe Arg Ser Val His Cys Leu Leu Tyr Pro Asp Thr Pro Trp 355 360 365

Cys Pro Trp Lys Ala Ala Tyr 370 375

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 29 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:	
CACCGTCGAC CCGCCGCC CTCCACTGC	29
(2) INFORMATION FOR SEQ ID NO:5:	
 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 37 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
(ii) MOLECULE TYPE: cDNA	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:	
GGTCTAAGCT TTGGCCATTA GAAGATGGCA GTGCGGG	37
(2) INFORMATION FOR SEQ ID NO:6:	
 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 34 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
(ii) MOLECULE TYPE: cDNA	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:	
GCGTCTAGAC CGCCATCATG CTCAAGCGCT GCGG	34
(2) INFORMATION FOR SEQ ID NO:7:	
 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 22 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
(ii) MOLECULE TYPE: cDNA	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:	
GTAATACGAC TCACTATAGG GC	22

(2) INFORMATION FOR SEQ ID NO:8:

	(i)	SEQUENCE CHARACTERISTICS: (A) LENGTH: 32 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
	(ii)	MOLECULE TYPE: cDNA	
	(xi)	SEQUENCE DESCRIPTION: SEQ ID NO:8:	
CGT	CTAGA	CG CGGCCGCCAC CCCACCATGC TC	32
(2)	INFO	RMATION FOR SEQ ID NO:9:	
	(i)	SEQUENCE CHARACTERISTICS: (A) LENGTH: 29 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
	(ii)	MOLECULE TYPE: cDNA	
		SEQUENCE DESCRIPTION: SEQ ID NO:9:	
		GA CCATGGCCAC TAGAAGATG	29
(2)		RMATION FOR SEQ ID NO:10:	
	(i)	SEQUENCE CHARACTERISTICS: (A) LENGTH: 56 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
	(ii)	MOLECULE TYPE: cDNA	
	(xi)	SEQUENCE DESCRIPTION: SEQ ID NO:10:	
TGGC	GTCTA	GA CCATGGCCAC TAAGCGTAGT CTGGGACGTC GTATGGGTAG AAGATG	56
(2)	INFO	RMATION FOR SEQ ID NO:11:	
	(i)	SEQUENCE CHARACTERISTICS: (A) LENGTH: 418 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: both (D) TOPOLOGY: linear	

(ii) MOLECULE TYPE: cDNA

(xi) SE	EQUENCE DESC	CRIPTION: SE	EQ ID NO:11:	:		
TGGTTTGCNA	CGGGNGGCGC	TGGCTTCTNC	ATCANCCGTG	GGCTGGCTCT	AAAGATGAGC	60
CCGTGGGCCA	GCGGGGGTNA	CTTCATGAAT	ACGGCTGAGC	GGATCCGGCN	GCCTGATGAC	120
TGCACCATCG	GCTACATCGT	GGAGGCCCTG	CTGGGTGTGC	CCCTNATCCG	CAGCGGCCTN	180
TTCCACTCCC	ACCTGGAGAA	CCTGCAGCAG	GTGCCCACCT	CGGAGCTCCA	CGGAGCAGGT	240
NACGCTGAGC	TTACGGTATT	TTTTGAAAAA	AAGCGGGAAC	GCCGTTCCAC	GTNAAAGGGG	300
GCCTTTTTTN	GGTTGGGAGG	GCCGACCCAT	TCCAGGTTTC	CGTTTNCATT	CCANTTGCCA	360
ACTTTTAACC	NGGGAAAAA	CCTTGGTNTT	TCCCCGAAAT	TGNAATTTTT	TAGGTTGG	418

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 443 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: both
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GGCACGAGTG AAGGGCTGCC TGCTGAAGGC CGATTTTCTC CTTCCAGACG TTCATCTTCA 60 CTGACGGGGA AGATGAGGCC CTGGCCAGGC ACACGGGTGA GCCCTGGACT TGGGGCGGGA 120 NGGGGCCCAA GCCTCCATCC AGAGCCGAAC GCTCCCCTC CAGTCTCCCT GCCCCTCTGG 180 GCCGAGAAGT CACCAAGGGC AGGACANGGA GGGCAGTTTA CTCATGGGGT TTGCTCCATG 240 CCCCGGCCCA ACACTCGGGN CCCCCAATTC TCATGCAAAT GAAGCCCATT CAGCCCCCCG 300 GGTTCCTTTG AGCCAAGCAG CGGCAAAAGT GGCGGTTCTG GNAAAAGTGC TGATTGGCGG 360 GGCGGGCCC CTTGAAGGT TGGTTCGGAA GTCAAAGCCA AGCCCGAGTA AGTTGGGAGT 420 ANGCGCCGGC CAGTTAANAA GGG 443

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 513 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: both
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(>	(i) SE	EQUENCE DESC	CRIPTION: SE	EQ ID NO:13:	:		
CTCGAC	STCCG	GGTGCAGGTG	GCAATGGACA	GAGCGGAACC	TGGATGGGTC	AGCTTCCACA	6
GAGAAT	GGTC	CCTTGATGTG	CACTGCGTTC	CGCTTGTTCT	CAAACATGCC	ATAGCTCAGG	12
GTCACO	CTGCT	CATGAAGCTC	GGTGGTGGGC	ACCTGCTGCA	GGTTCTCTAG	GTGGGAGTGG	18
AAGAGG	BCCGC	TCCGGATGAG	GGGTACACCC	AGCAGAGCCT	CTACAATGTA	GCCAATGGTG	24
CAGTC	ATCGG	GGAGCCGGAT	GCGCTCTGCC	GTGCTCATGA	AGTGTCCTCC	ACTGGCCCAT	30
GGGCC	ATCTT	TAGGGCCAGC	CCTCGGCTGA	TGCAGAAGCC	AGCTCCTCCG	GTGGCAAACC	36
AAAAGT	GGAC	AGGTCTCACT	TTGTGCTCGC	TGATCCGTTC	TGTGGCCTGG	ATGGGCCTGT	42
CCAGGC	CTGGG	CTTGCGATGT	ACACGTCTTG	GGTGTGGGGA	TAGCTGGCCA	GGAAGCGCAG	480
CAGCGC	CCGG	AAGTTGACGT	AGTTGTCATC	ATC			513

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 352 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: both
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GCGCCCCTG GCCGAGCCGC TNCGCGCCCC GAGACGTCTT CATCGCTGTC AAGACCACCA 60

AAAAGTTCCA CCGCGCGCC CTCGACCTGC TGCTGGAGAC CTGGATCTCG CGCCACAAGG 120

AGATGACGTT CATCTTCACT GACGGGGAAG ATGAGGCCCT GGCAGGCACA CGGGCAACGT 180

GGTCATCACA AACTGCTCGG CCGCCCACAG CCGCCAGCTT GNTTCCTGCA AGATGGCCGT 240

GGAGTATGAC CGCTTCATCG AGTCCGGCAG AAGTGGTTCT GCCACGTGGA CGATGACAAC 300

TACGTCAACC TGCGGGCCCT GCTNCGGCTG CTGGCCAGCT GACCCGCACA CG 352

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 449 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: both
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

ACAGGTCTCA	CTTTGTGCTC	GCTGATCCGT	TCTGTGGCCT	GGATGGGCTG	TCCAGCTGGG	60
CTTGCGATGT	CACGTCTTGG	GTGTGGGGAT	AGCTGGCCAG	GAGCCGCAGC	AGCGCCCGGA	120
GTTGACGTAG	TTGTCATCAT	CCACGTGGCA	GAACCACTTC	TTCCCAGACT	CAATGAATCG	180
GTCATACTCC	ACAGCCATCT	TGCAGGACAG	AGCCTGGCGC	TGTGGGCCGA	GGAGCAGTTG	240
GTGAGCACCA	CATTGCCTGT	GAGCTTGGCC	AGAGCTTCGT	CCTCCCCATC	AGTGAAGATG	300
AACGTCATCT	CCTTGTGGCG	CGAGATCCAG	GTCTCGAACA	GCAGATCGAG	CCGCGCGCGG	360
TGAAACTTTC	TGGTGGTCTT	GACGGCGATG	AAGACGTCGC	GAGGGGACAG	AACTTTCGGC	420
GGGGGACGCG	GATGGGCGTC	GCCCTGCGA				449